

R E P O R T R E S U M E S

ED 011 128

EA 000 157

THE STUDY OF FACTORS INFLUENCING COLLEGE GRADING STANDARDS.

BY- HILLS, JOHN R. GLADNEY, MARILYN B.

UNIVERSITY SYSTEM OF GEORGIA, ATLANTA

REPORT NUMBER R-BULL-2-66

PUB DATE 25 APR 66

EDRS PRICE MF-\$0.09 HC-\$0.72 18F.

DESCRIPTORS- *COLLEGES, *ACADEMIC APTITUDE, *ADMISSION CRITERIA, *DORMITORIES, *GRADING, SURVEYS, COLLEGE STUDENTS, STUDENT ENROLLMENT, ADAPTATION LEVEL THEORY, ATLANTA

THE STUDY TESTED THE HYPOTHESIS THAT COLLEGE STUDENTS WITH A SPECIFIED AVERAGE ACADEMIC APTITUDE (AAA) LEVEL RECEIVED HIGHER GRADES AS ENROLLMENT INCREASED AND ADMISSION STANDARDS WERE LOWERED TO FILL NEW DORMITORIES. AAA WAS COMPUTED FROM AN ENTERING FRESHMAN'S COMBINED SCHOLASTIC APTITUDE TEST (SAT) VERBAL AND QUANTITATIVE SCORES. LOCATING COLLEGES WHICH MET THE CONDITIONS OF THE HYPOTHESIS WAS DIFFICULT. IN THE 140 COLLEGES SURVEYED, ENROLLMENT TENDED TO INCREASE AT ABOUT THE SAME RATE AS ENTERING FRESHMEN AAA--NEW DORMITORY SPACE WAS USED TO REDUCE CROWDING, TO REPLACE OBSOLETE QUARTERS, AND TO HOUSE STUDENTS OTHER THAN FRESHMEN. IN THE SIX COLLEGES USING NEW DORMITORIES FOR FRESHMEN, THE AAA INCREASED IN TWO, REMAINED THE SAME IN TWO, AND DECREASED IN TWO. THE DATA FROM THE LAST TWO COLLEGES, HOWEVER FAILED EITHER TO SUPPORT OR REJECT THE HYPOTHESIS. THE HYPOTHESIS, THEREFORE, WAS REVISED TO EXCLUDE THE HOUSING AND ENROLLMENT VARIABLES. USING THE GULLIKSEN-WILKS ANALYSIS TO TEST THE REVISED HYPOTHESIS ON THE DATA FROM FIVE COLLEGES, THE FINDINGS SUGGESTED THAT (1) GRADING STANDARDS TENDED TO BECOME MORE SEVERE AS AAA INCREASED, (2) ISOLATED AAA INCREASES MIGHT NOT HAVE RESULTED IN OBSERVABLE GRADING STANDARD CHANGES, (3) GRADING STANDARD CHANGES LAGGED BEHIND AAA CHANGES BY 1 YEAR. (JB)

ED011128

Research Bulletin 2-66

THE STUDY OF FACTORS INFLUENCING COLLEGE GRADING STANDARDS

John R. Hills

and

Marilyn B. Gladney

Office of Testing and Guidance
Board of Regents
University System of Georgia
244 Washington Street, S.W.
Atlanta, Georgia 30334

April 25, 1966

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE
PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION
POSITION OR POLICY.

EA 000 157

Research Bulletin 2-66

The Study of Factors Influencing College Grading Standards

by

John R. Hills
and
Marilyn B. Gladney

It has been observed several times that the standards associated with various levels of course grades (A, B, C, etc.) in colleges may vary over time (Aiken, 1963; Hills, 1964; Hills and Klock, 1965; Webb, 1959). Usually what has been found is that from year to year admissions standards have been raised at these colleges, most often as a result of increased numbers of applications for the available classroom or dormitory spaces. On one occasion (Hills, 1965 mimeo), it was observed that one institution whose admissions standards had been rising, abruptly broke that trend during one year when a new dormitory was opened. Upon lowering admissions standards to fill the new dormitory, grading standards sank. It became easier temporarily to obtain good grades. The following year the trend toward raised admissions standards--and grading standards--resumed.

Adaptation Level

The phenomenon observed in these studies falls into the class of phenomena which are explained by Helson's theory of adaptation level (Helson, 1947, 1948, 1951). This theory states that experience is ordered and that one of the chief characteristics of its patterning is the establishment of neutral, indifferent, or zero regions of functioning. Such points or regions represent the adaptation-level of the organism with respect to given stimuli. The adaptation-level tends to be determined largely by the values of the background stimuli and the residual from stimuli in past experience. All effects of stimulation, past as well as present, are pooled to form a single level with respect to given classes of stimuli, but the effects of past experience are usually not sufficient to displace the adaptation-level which develops from within the immediate-stimulus continuum. The adaptation-level is a moving, changing region, and it may change even in the absence of stimulation. Such a concept accounts both for invariants in perception and judgment and also for individual differences in perception and evaluation in the face of the same stimuli.

The translation of college grading practices into the adaptation-level model is direct and straightforward. The stimuli are the performances of students. A stimulus to be judged is the performance of a single student. The background stimuli are the performances of other students in the same class. The residual stimuli include the performances of students in other

similar classes in the faculty member's past-experience, the performances of students in different classes, the faculty member's own performances in similar situations, etc. According to the theory, all these stimuli are pooled to determine the faculty member's neutral point in grading a student's performance. However, the residuals from past experience are not sufficient to overcome the effects from the students in the immediate class. Thus, if the level of performance of the immediate class is higher than that of previous classes in the professor's experience, the adaptation-level, i.e., the level of performance which the professor thinks of as neither particularly good nor particularly bad, shifts upward for individual performances of members of this more proficient class. This explains why Aiken, Webb, and Hills found that as admissions standards rose, the average college grade did not rise proportionately.

Factors Influencing the Level of Background Stimuli

The background and residual stimuli for college grading are the performances of other members of the immediate class and the performances of members of other classes. Typically the performance of members of college classes is associated with the academic potential of the members of those classes. For this purpose, academic potential can be defined as level of performance on admissions tests and in previous academic situations, i.e., high school grades. Abundant evidence supports the statement that for groups of students there usually is a substantial multiple correlation between academic potential, thus defined, and performance defined by average college grades. Thus a factor which would influence the level of background stimuli in the classroom situation would be the level of academic potential of the members of the class.

Academic potential of classes is largely determined by the college admissions and recruiting functions. For example, if recruiting is effective, a college's applicant pool is likely to increase in size. If there is no increase in academic facilities, or the increase in facilities is not proportional to the increase in the applicant pool, then a smaller proportion of the applicants can be admitted. If applicants are admitted primarily on the basis of their academic potential, the result would be an entering class of greater academic potential than previous classes. This, in turn, provides a new set of background stimuli at a higher level than the residual from past experience, and according to Helson's theory the effect from past experience is not sufficient to keep the adaptation-level from shifting upward.

If new dormitories are opened providing additional space for freshmen in the face of an applicant pool which has not increased proportionately, then a higher proportion of applicants can be admitted. Again, if admission is based on academic potential the admissions office dips deeper into the pool, admitting students with less potential than the minimum of former classes, and adaptation-level shifts downward. A less competent performance is at the neutral point, and grading has become easier. These two factors, the amount of facilities and the size of the applicant pool, seem likely to be the most significant factors which will influence the level of background stimuli for college grades.

There could be a functional relationship between change in the proportion of applicants admitted and grading standards (adaptation-level). One might expect to find that, all other things being constant, the higher the proportion admitted, the lower the adaptation level, assuming that admission is based on academic potential. However, in the practical situation there are many other things which refuse to remain constant. Increase of dormitory facilities, for instance, does not always mean that larger numbers of freshmen can or will be admitted. Sometimes dormitories provide space needed for upper classmen. Dormitories are sometimes shifted from use by one sex to use by the other sex. Some admissions officers carefully manipulate between campus housing and off-campus housing so that when a new dormitory opens no appreciable change in admissions takes place. The dormitory space is filled by people who formerly were attending the institution but living off the campus. Sometimes off-campus housing is deliberately built, i.e., private dormitories, not under control of the college. These can sometimes make it possible for an institution to increase its enrollment without any change in the institution's facilities. If such a dormitory is closed, the capacity of the institution might be reduced through no control or fault of its own, also. Sometimes population pressure can increase so rapidly that even the addition of new facilities does not let the institution maintain the same proportion of admissions to applications. This may be due to more effective recruiting, to changed birth rates eighteen years previously, to changed attitudes toward higher education, or to a complex mixture of these.

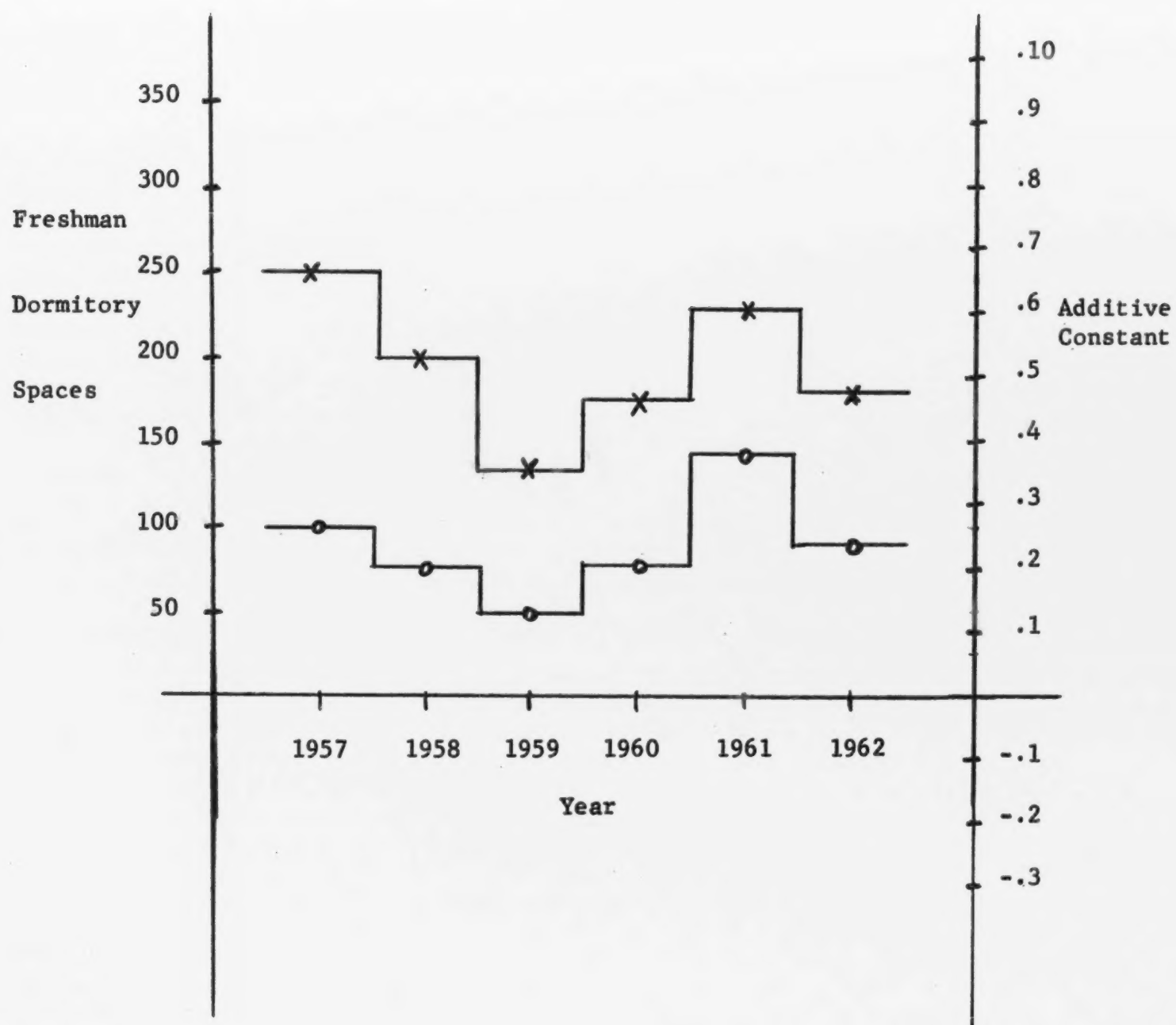
Techniques of Measuring Grading Standards

To study influences on adaptation level in the college grading situation one needs to be able to evaluate grading standards. Several procedures are possible, but one which has seemed quite effective to the authors is a variety of analysis of covariance (Gulliksen and Wilks, 1950). For the purposes of this kind of study, one may think of the Gulliksen-Wilks analysis as revealing (providing the assumptions are met) whether for the same academic potential similar college grades are received by different groups of students. For our purposes, academic potential is measured by the combination of SAT scores and high-school average grades which is optimally related to college grades. Upon discovering such a combination which fits all of the groups of students to be compared, one evaluates grading standards (adaptation-level) by examining the additive constant appropriate for each group to make most accurate estimates of the actual grades obtained by members of that group. If the additive constant becomes less positive (or more negative), grading standards have become more severe. The standard, or the adaptation-level, has gone up. If the constant becomes more positive (or less negative), more is being added in for each student, implying that grading is less severe, and standards or adaptation-level have gone down.

To see whether adaptation-level has followed admissions, academic potential changes, changes in housing or other facilities, etc., graphs showing both functions over time can be drawn, as in Figure 1. The base line in that illustration, which is idealized, not based on real data, is a scale of time--successive fall entering freshman classes. The ordinate on the left is a

Figure 1

Illustration of Graph of Standards (Constants) and Academic Potential



scale of dormitory spaces for freshmen (the "x's" in the graph). To have the kind of fluctuations which appear in this illustration a college might have been embarking on a building program which required that it demolish old dormitories in order to build new ones on the same ground. Thus the number of spaces decreased until 1960 when the first new dormitory was completed. Then, one might conjecture that one of the new dormitories was destroyed or otherwise became unavailable for freshmen. The scale on the right, associated with the "o's" in the graph, is the scale of grading standards or adaptation-level expressed in terms of additive constants. This illustration has been drawn to show what such a graph would look like if grading standards fluctuated in unison with dormitory spaces. The lines of values parallel each other. (The lines have been drawn as step functions since such changes tend to occur on an annual basis rather than gradually during a year.) As the number of dormitory spaces goes down between 1957 and 1958, the additive constant becomes less positive, i.e., lower grades are obtained for a given aptitude. As the number of spaces goes up between 1960 and 1961, the constant becomes more positive. Grading standards are eased. Presumably the basis for this parallelism would be the fact that as the number of spaces decreased, smaller proportions of applicants could be admitted. Since admission was based on academic potential, those admitted were more able. But those more-able students also formed a set of immediate background stimuli which overpowered the faculty members' residual stimuli from past years, and adaptation-level went up. In 1960 and 1961 adaptation-level came back down. If the theory does not hold for this kind of stimuli, then we would not expect the lines to be parallel. However, lack of parallelism would not be expected if extraneous variables interfered with the data. For instance, if the birthrate in 1941 was less than in 1940 then the applicant pool might have decreased about as much as the dormitory spaces decreased. Then it might be that no change in academic potential or in the proportion of applicants admitted would take place, and no parallel change in additive constant could be expected.

Data

In the University System of Georgia, a complex of public higher education facilities involving numerous colleges, during the period from 1960 through 1964 there were nineteen units admitting freshmen each fall. Three of these admitted only men or women. These institutions provided thirty-five sets of entering freshmen data for students of like sex for four pairs of adjacent years, making a total of 140 situations in which changes could be observed in grading standards related to changes in admissions standards, which might be further related to changes in housing facilities on campus. During this period we were able to verify the opening of 33 new dormitories or annexes to dormitories among these colleges. Some were very large; one was big enough to accommodate 1000 students in a single building. Others were rather small. Two of the institutions opened as many as five new dormitories in a single year. These were the largest institutions in the System. The size of institutions in the System varies from under 1000 to over 10,000 in total enrollment.

An initial tabulation of changes in freshman enrollment from year to year revealed that numerous changes of large size, on the basis of percentage of entering freshmen, had taken place during this period. At some institutions there were increases of over 100% in the number of entering freshmen

men or women from one year to the next. There was a distinct tendency for enrollment to increase rather than decrease during this period, though there were some sizable decreases.

Finally, there were some substantial increases from one year to the next in average academic potential of entering freshmen during this period. For convenience, the average academic potential for this purpose was measured by the sum of the average SAT Verbal and Mathematics scores. The largest change between adjacent years was 93 points for males at one of the four-year colleges. There were five changes in means greater than 63 points, and 11 greater than 50 points.

When one seeks to find the situations which provide test cases for the complete hypothesis that grading standards adapt to housing facilities via the midwife of admissions standards, much of this potential disappears. Many of the new dormitories did not provide any additional housing for freshmen. Some did not provide additional housing at all, but merely replaced obsolete quarters. Others were not made available to freshmen and did not release other space for freshmen but merely provided quarters for other kinds of students, such as married couples, upperclassmen, etc. Some did not provide facilities that permitted admission of more freshmen, but merely alleviated the serious overcrowding (three and four students housed in rooms designed for two) that had been in existence. Of the 140 possibilities for occurrence jointly of large proportional increases in entering freshmen and opening of new dormitory facilities which increased the space for freshmen, only six actually occurred during this five-year period.

The third concomitant required for full test of the hypothesis is that along with the joint occurrence of new dormitory space for freshmen and increased numbers of freshmen is the decrease in average academic aptitude of the entering group. In the case of two of the six possible situations, with new space and additional students, the average academic aptitude actually increased. Probably the increase in size of the applicant pool was even greater than the proportional increase of admissions--the years were 1963 and 1964 which were population-boom years. In two more of the situations the academic aptitude did not change, the change in sum of V+M being 5 points or less. This leaves us with two situations in which the complete hypothesis can be evaluated. In one of these, Institution A, 64 new spaces were provided for freshman females, 33 more females enrolled in the freshman class, and academic aptitude (V+M) decreased by 37 points. In the other, Institution B, 150 new spaces were provided for freshman males, enrollment increased by 100, and the sum of V and M decreased by 16 points. Neither of these decreases in aptitude is very large from a practical standpoint, but from these five years of data no other likely situations emerge. Probably the expansion currently taking place in higher education enrollments makes this an inauspicious period in history for the testing of the complete hypothesis.

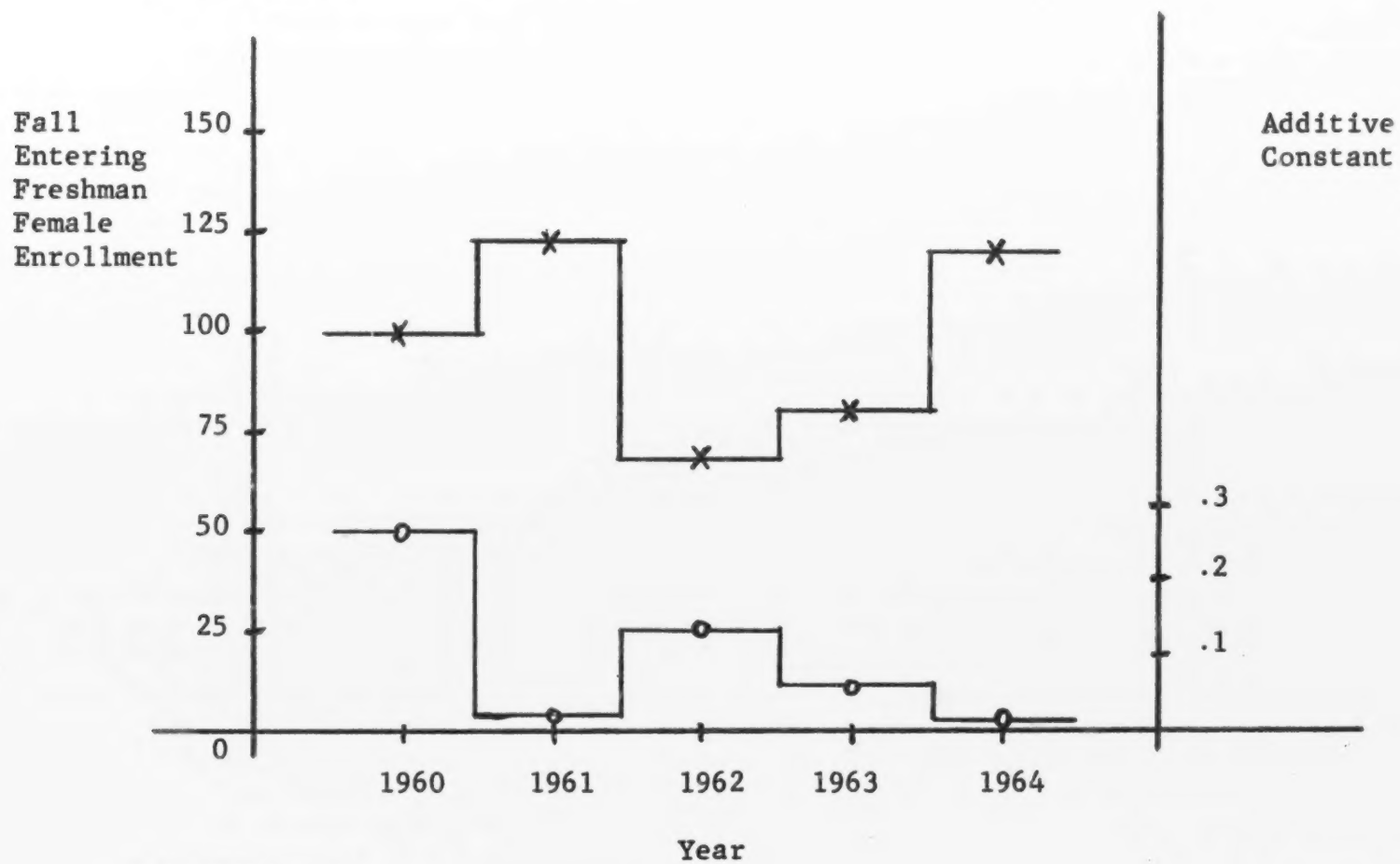
Results

As can be seen in Figure 2, for Institution A the lines on the graph are anything but parallel, while for Institution B the lines are more nearly parallel. Thus, in one instance the hypothesis in its complete form is not

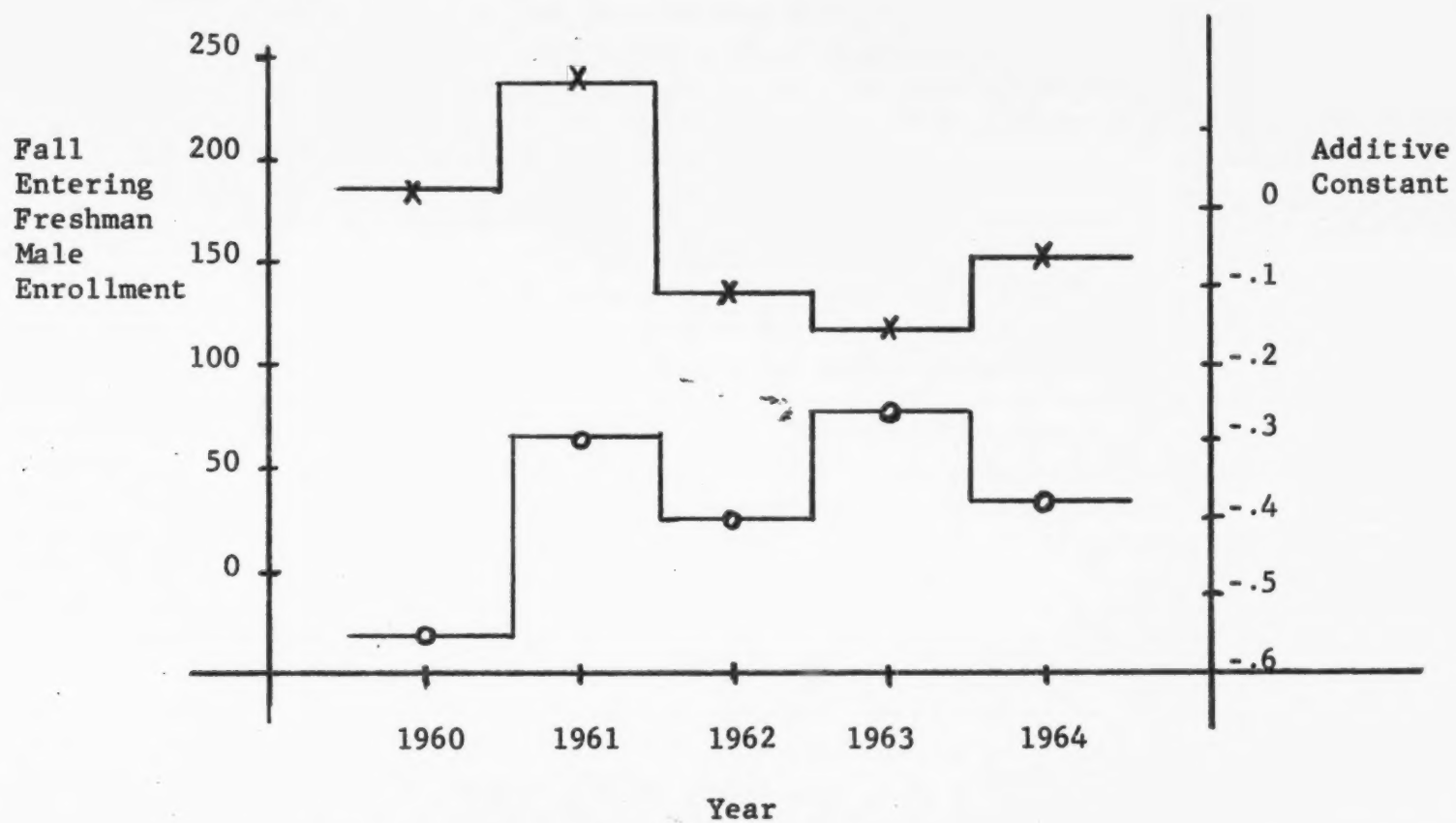
Figure II

Enrollment vs. Grading Standards

Institution A



Institution B



well supported while in the other it is only modestly supported. Such a result is inconclusive, especially when it is recognized that Institution B data involving the years in question were substantially involved in the formulation and initial elaboration of the hypothesis. They partially overlapped the data reported in Hills' mimeographed 1965 report, which showed a clear parallelism of such data at Institution B.

Further Considerations

The basic problem with the complete hypothesis at this point in the history of higher education is that the times do not seem to be ripe for situations to occur in which it can be evaluated. What would happen if we looked only at part of the hypothesis? We could relax the restriction that changes in grading standards had to be associated with changes in housing in order to be of interest. And we could relax the requirement that to be worthy of study changes in admissions standards had to be concomitant with large proportional increases or decreases in enrollment. In other words, let us examine the more limited hypothesis that large changes in academic aptitude (V+M scores), for whatever reason, are associated with changes in grading standards. This is still a hypothesis within the framework of adaptation-level theory.

Returning to our 140 sets of data over the five-year period from 1960 through 1964, there are those five sets of data which include adjacent years during which occurred increases in average aptitude of more than 63 points on V+M. Are those increases in aptitude associated with decreases in grading standards (less-positive additive constants)? For this hypothesis to be supported we should expect the lines on the graph to diverge. If they are parallel, the hypothesis is contradicted. The difference is that earlier we were graphing increases in enrollment with the implication that increases in enrollment implied decreases in aptitude. Now we are graphing the changes in aptitude, themselves.

Further Results

The graphs for these five sets of data appear in Figures 3 through 7. Again notice that the scale on the left of the graphs goes with the line of x's, while the scale on the right side accompanies the line of o's. The x's represent academic aptitude as measured by the sum of the average SAT V and M scores each year. The o's represent grading standards in terms of the additive constant for a common regression equation. If the additive constant becomes less positive (or more negative), grading has become more severe because grading standards have been raised. All five of these graphs have a general tendency for the lines to diverge, with the tendency becoming less clear as we proceed from Figure 3 to Figure 7. In Figure 7 we have the picture of aptitude taking a single jump during these five years and grading standards fluctuating up and down during the period. Generally the set of graphs would have to be taken as support for the hypothesis that as aptitude increases appreciably, grading standards tend to become more severe. They also suggest that isolated increases in aptitude may not result in clearly observable changes in grading standards. Since we only examined the most

Figure 3

Institution C: Aptitude vs. Grading Standards

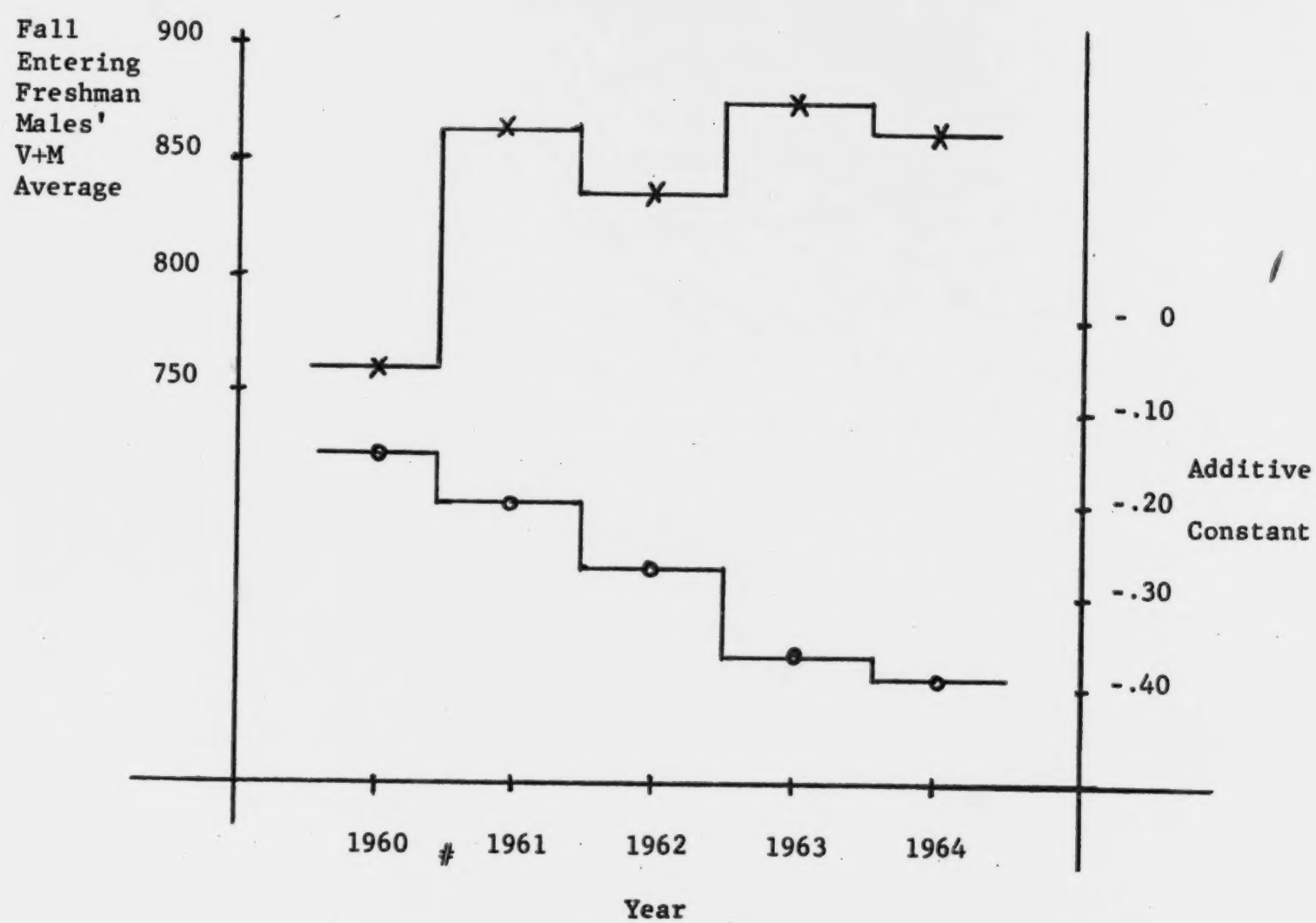


Figure 4

Institution A: Aptitude vs. Grading Standards

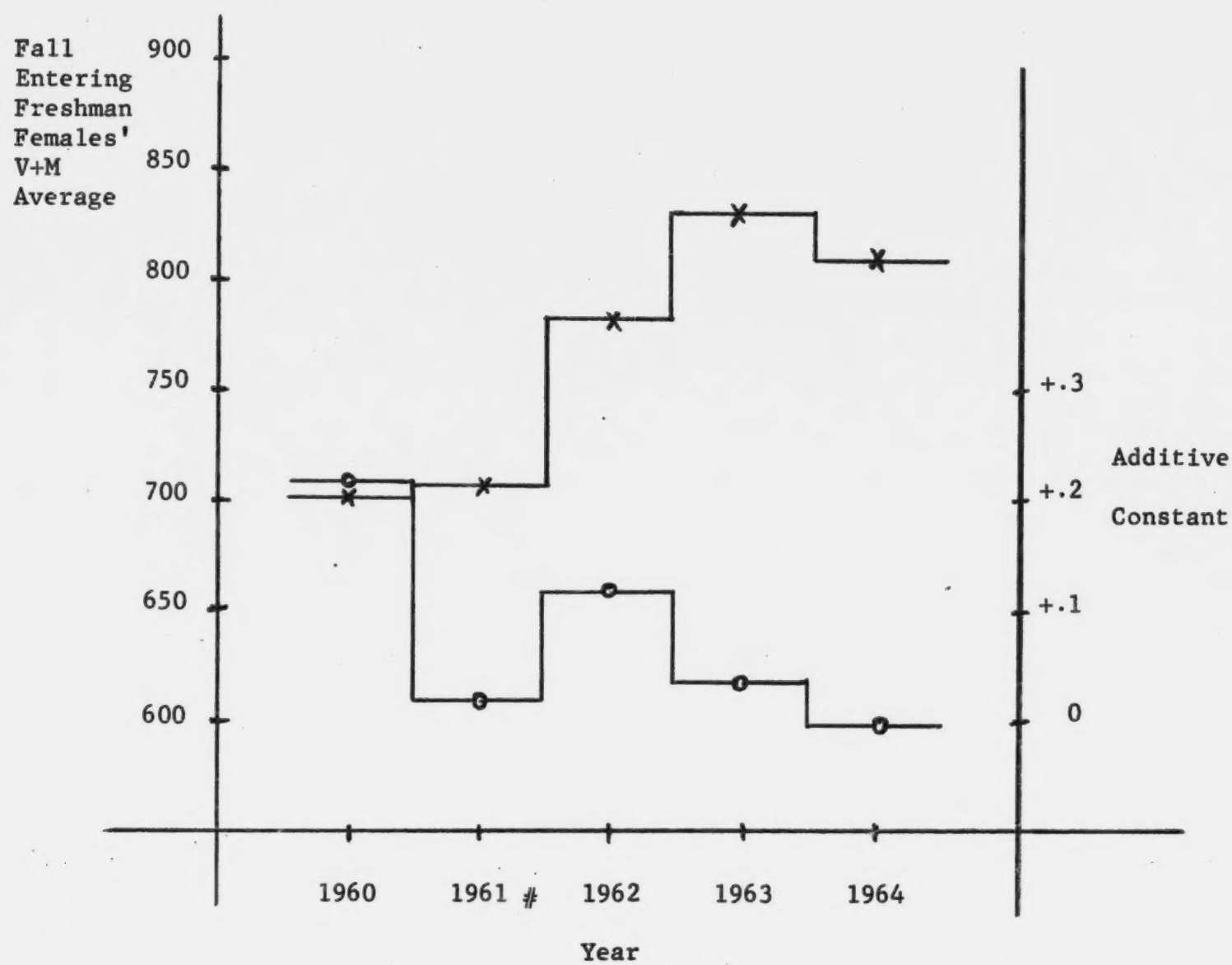
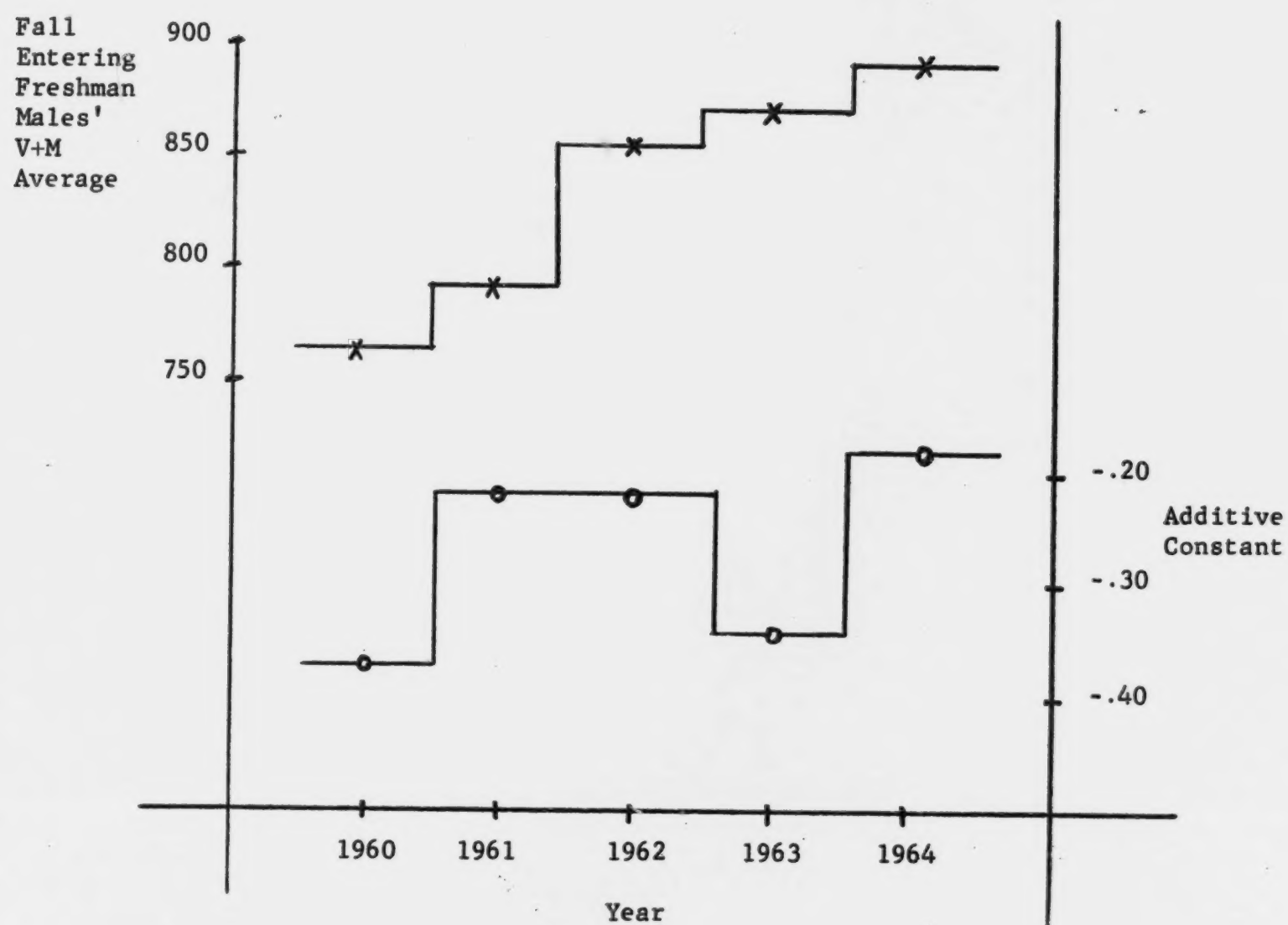


Figure 5

Institution D: Aptitude vs. Grading Standards



x= Aptitude
o= Grading Standard

Figure 6

Institution E: Aptitude vs. Grading Standards

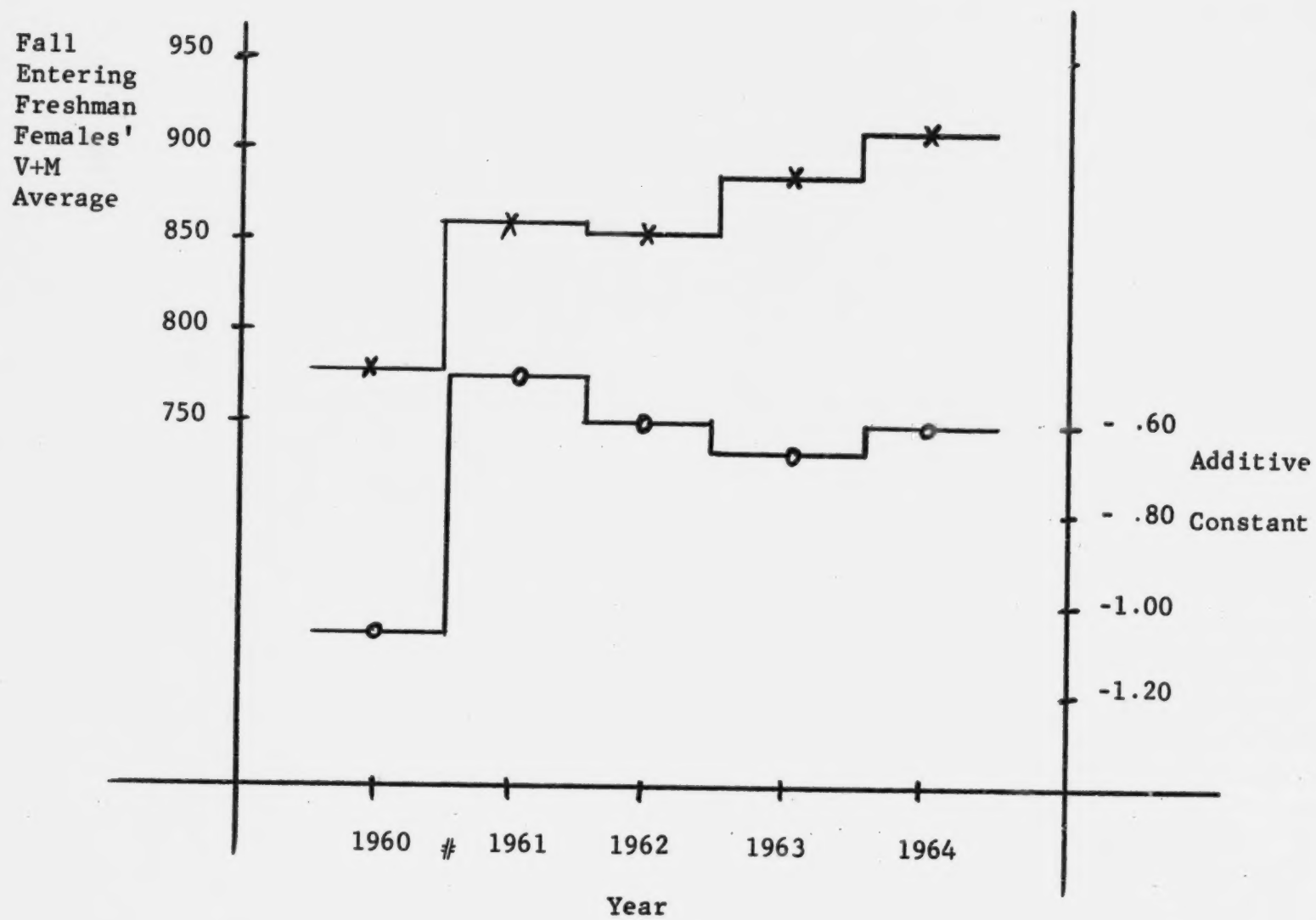
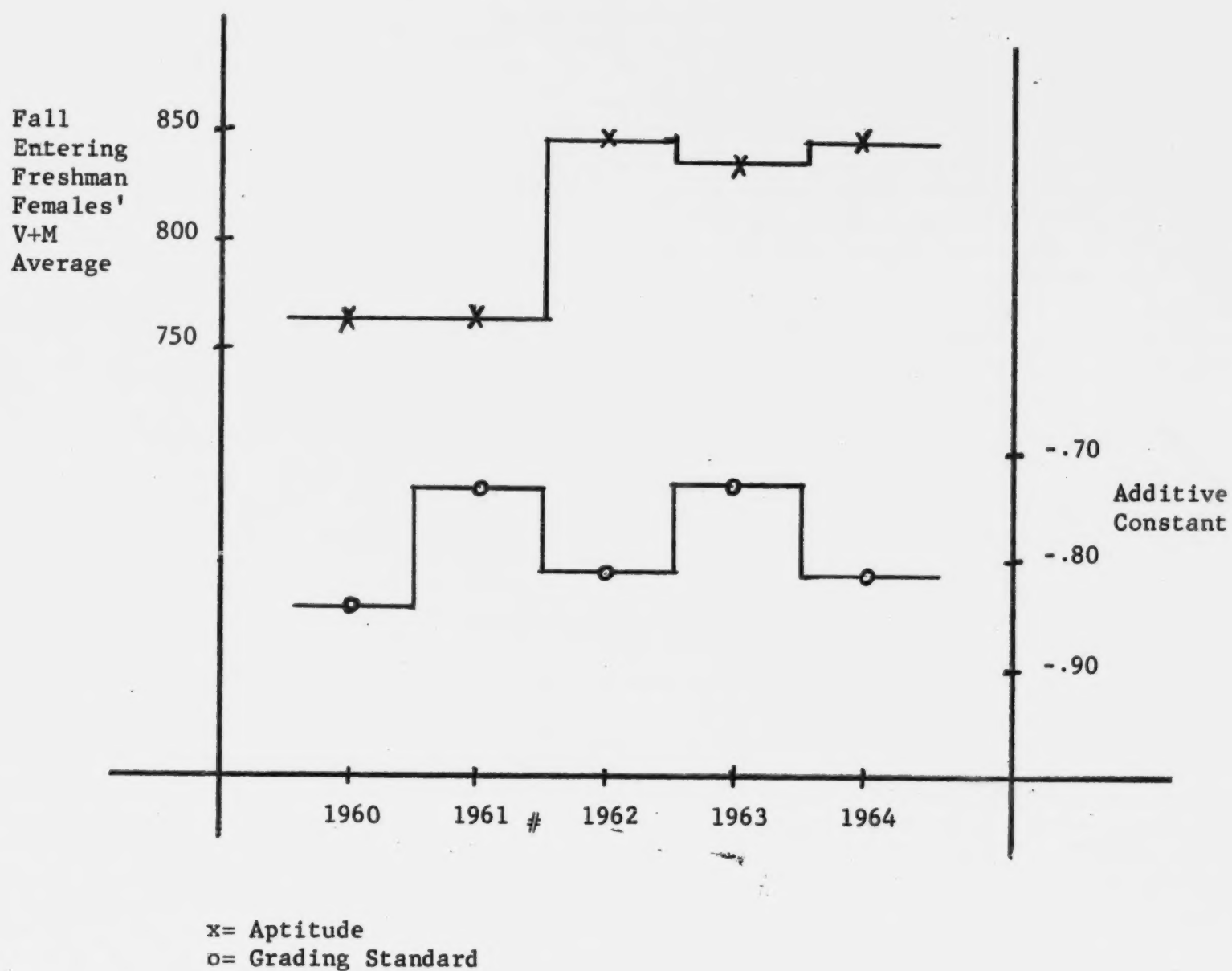


Figure 7

Institution F: Aptitude vs. Grading Standards



extreme annual increases in aptitude available in these data, there is implicit in our procedure the assumption that modest increases in aptitude will not produce very noticeable changes in grading standards.

There is another hypothesis suggested by these graphs. Notice on each of them there is a # mark between two of the years on the base line. This identifies the adjacent years between which the big increase in aptitude took place--the change which caused these situations to be singled out for study. If a full description of the phenomena was provided by our hypothesis that grading standards were determined solely by the adaptation level due to background stimuli, then we should expect the largest changes in grading standards to occur between the two years contiguous with the # on the base line. However, in four of the five graphs, Figures 3 through 6, a larger change in grading standards occurs one year later than the change which occurs on those contiguous years. This suggests that faculty may attempt to compensate in their grading practices when they know that a new, more able, group has entered, and that the attempt to compensate wears off after a year, the faculty going back to their old practices of centering their grade distributions on C regardless of the aptitude of the students. Testing of this hypothesis will have to await the collection of new data.

Summary

Some of the factors which might influence college grading standards were studied in the framework of adaptation-level theory. It was hypothesized that as dormitories provide additional space for freshmen students causing admissions standards to be lowered, grading standards become less severe. That is, higher grades are given students of a specified aptitude level. Out of 140 possible situations in which the conditions for a test of this hypothesis might have been present, only two possessed the necessary elements. They did not provide clear support for the hypothesis or clear negation of it. The hypothesis was relaxed to permit more use of the data. Five sets of data showed large increases in aptitude between adjacent years. These sets of data supported the hypothesis that grading standards became more severe as academic aptitude increased (regardless of factors which might have produced the increase), but they suggested that the change in grading standards might not be immediate. The faculty might compensate for the raised aptitude by giving higher average grades during the year that aptitude is first increased, but the data suggest that standards a year later become more severe as the faculty drift back to their customary proportions of C's.

References

- Aiken, Lewis R., Jr. "The Grading Behavior of a College Faculty." Educational and Psychological Measurement, 1963, 23, 319-322.
- Gulliksen, H. & Wilks, S.S. Regression tests for several samples. Psychometrika, 1950, 15, 91-114.
- Helson, H. "Adaptation-level as frame of reference for prediction of psychophysical data." American Journal of Psychology, 1947, 60, 1-29.
- Helson, H. "Adaptation-level as a basis for a quantitative theory of frames of reference." Psychological Review, 1948, 55, 297-313.
- Helson, H. (Ed.) Theoretical Foundations of Psychology. New York: D. Van Nostrand, 1951, pages 379-385.
- Hills, John R. "The Effect of Admissions Policy on College Grading Standards." Journal of Educational Measurement, 1964, 2, 115-118.
- Hills, John R. "Housing and Grading Practices." Research Bulletin 65-3, Office of Testing and Guidance, Board of Regents of the University System of Georgia. 1965, mimeo.
- Hills, J.R. & Klock, J.A. "Predicting Grades for all Schools in a Large University." American Educational Research Journal, 1965, 2, 145-150.
- Webb, Sam C. "Measured Changes in College Grading Standards." College Board Review, 1959, No. 39, 27-30.